

## Durham Research Online

---

### Deposited in DRO:

08 November 2010

### Version of attached file:

Published Version

### Peer-review status of attached file:

Peer-reviewed

### Citation for published item:

Song, W. and Zhou, Z. (2006) 'A contextual based semantic modeling approach to task-service formation in virtual organization.', in IEEE/WIC/ACM International Conference on Web Intelligence, WI 200, 18-22 December 2006, Hong Kong ; proceedings. Los Alamitos, CA: IEEE, pp. 653-656.

### Further information on publisher's website:

<http://dx.doi.org/10.1109/WI.2006.8>

### Publisher's copyright statement:

© 2006 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

### Additional information:

## Use policy

---

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in DRO
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full DRO policy](#) for further details.

# A Contextual Based Semantic Modeling Approach to Task-Service Formation in Virtual Organization

William Song<sup>1\*</sup> and Mingquan Zhou<sup>2</sup>

<sup>1</sup> Computer Science, Durham University, South Road, DH1 3LE, England, UK

<sup>2</sup> Computer Science, Beijing Normal University, Beijing, China

w.w.song@durham.ac.uk

## Abstract

*Virtual organizations are considered to be an independent mechanism, which manages to bridge the users' goals and requirements to the grid/web resources and services. To form the workflow for a virtual organization we need to find a sequence of interrelated services (the grid/web resources) matched to given users' requirements. It is crucial to find a semantic description for virtual organizations in order to analyze various components, such as tasks, services, and resources, and hence to make virtual organization workflows through semantic matching between tasks and services. In this paper, we propose a contextual based semantic description approach to the semantic description of virtual organization components, tasks and services. The contextual information for a resource or a service or a task is the information provided in the application domains and the pre-defined (standardized) service ontology descriptions. We also propose a semantic matching theory for matching the tasks with the services.*

## 1. Introduction

The grid is an emerging platform to support on-demand "virtual organizations" for coordinated resource sharing and problem solving on a global scale [6]. The real and specific problem that underlies the grid concept is coordinated resource sharing and problem solving in dynamic, multi-institutional virtual organizations [4]. A virtual organization is formed to temporarily and dynamically manage and organize a collection of resources or services (sometimes called nodes) for a given purpose or task. Each node in the organization plays a different role and performs a different activity and these nodes should coordinate and cooperate so that the different efforts and functions can be integrated to achieve a given goal.

Because these nodes in the grid/web are usually heterogeneous, autonomous, have different structures and purposes, and possess different characteristics and ontologies, it is unavoidable for the nodes to experience a painful process of communications and understanding before reaching a consensus and forming a virtual organization. During this process, commonalities and distinctions in perceptions of organizational structures, domain problems, community terminologies, business strategies, and domain ontologies are analyzed, described, and represented.

To build up a virtual organization we need to define a high-level task as a goal of the organization, for example, "to acquire the knowledge of Semantic Web". It is natural that this task will be decomposed into a number of subtasks. Each non-decomposable, leaf task corresponds to one or several services. The process of task decomposition supports forming the workflow of the virtual organization. The activities within a task include the process of requesting, comparing, selecting, consuming, integrating, and releasing services and resources.

Services are considered to be applications or functions used by the agents or other services. Services are distributed without any central control. Many services are orchestrated together to meet the requirements of a task. In the grid/web environment, there are a lot of resources and it is the function of services to organize and consume them.

Traditional workflow model concerns the execution dependence between tasks [16]. It also concerns the dependence between services available for the virtual organization and time sequence dependence.

The rest of the paper is organized as follows. We discuss the tasks and services and then introduce a context based semantic model for virtual organizations in section 2. In section 3, we discuss the semantic comparison and matches. In section 4, we conclude the paper.

---

\* Corresponding author: Computer Science, University of Durham, South Road, DH1 3LE, Durham, UK, Tel. +44-191 334 1729, Fax. +44-191 334 1701, Email. w.w.song@dur.ac.uk.

## 2. Semantic model for VO components

The core components in a virtual organization are tasks and services. Tasks are considered to be the final representation of users' demands and requirements whereas services are the conceptual clustering of resources. Between tasks and services, a semantic matching process is, based on the properties of the tasks, to discover the most suitable services that satisfy the tasks. Once this semantic match is completed, the tasks and the services will bind together to form the virtual organization workflow for the given objectives.

In this section, we will first discuss general descriptive (and contextual) information for virtual organization components and then propose the contextual based semantic characteristic model.

### 2.1. Tasks and services

In a virtual organization, tasks are the major components, which realize the end users requirements, construct workflows, and discover and match services. In other words, the tasks represent and formulate the users' demands and requirements at one end and deliver the basic services and consume the resources at the other end. It is obvious that the decomposition operation is a most fundamental one on the tasks. By applying the decomposition operation, the tasks form a tree structure, which supports to construct and schedule workflows for the virtual organization. A task is semantically described as follows:

- **Task name:** a concept name is used to represent the concept of a task.
- **Description:** a natural language description of the task.
- **Inputs:** a set of inputs to the task and their types. The inputs are from other tasks.
- **Output:** a set of outputs from the task and their types. The output is to an adjacent task.
- **Decomposition:** this is a specific relationship between two tasks. This relationship results in a set of subtasks. In OWL-S [11], this relationship can be viewed to be similar to `subClassOf`, `partOf`. This forms an ontology tree for the concept of the tasks and services.

Services are considered to be pieces of software or software components providing functions for the users to support their applications. The main problems with service are, among others, semantic description and discovery for services. OWL-S provides an ontology language for web services.

Our representation model considers the description capability for both the features at the high level of

business objectives for tasks and at the low, concrete level of functionalities for services. Therefore, using part of OWL-S capacity, a service is described as follows:

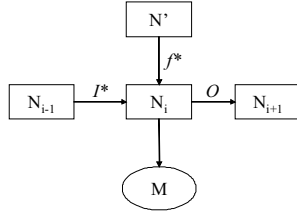
- **Service name:** the name for the service.
- **Description:** a set of data that describe the service, e.g., service description in natural language, service creator, service creation time, etc.
- **Inputs:** a set of inputs and their types. The inputs to the service are from other services. The types of the inputs can help restricting the selection of services.
- **Output:** an output and its type. The output of the service is usually an input of the other service. The type also helps with the selection of services.
- **Ontology:** for a given domain a group of ontologies have been pre-defined for describing services. It is important to note the group of domain related ontologies is critical for semantic matching between tasks and services.

### 2.2. Contextual based semantic model

The initial concept for semantic contextual model is based on the belief that the identification of meaning of a concept mainly stems from its contexts, i.e. its relationships to other concepts. For example, we can identify that A is an airline company in the sentence "company A provides services of transporting people and goods by air". In this example, only one pair <relationship, concept> (e.g. <transport, people>) for figuring out the meaning of A is, at most of time, not sufficient. Actually, in a given circumstance, we can manage to collect a number of such pairs for identifying a concept.

Using the contextual based semantic model proposed for tasks and services, we can construct a semantic description structure for a task and a service. For simplicity and without loss of generality, we will not distinguish the descriptive differences between tasks and services. We use node to represent a task or a service. A semantic description structure for a node is illustrated in Fig. 1.

**Definition 1 (Basic semantic characteristic model)** A semantic context model is a directed graph  $G = (V, E)$ , where  $V$  is a set of concepts and  $E$  a set of semantic relations between concepts. For each concept  $v$  in  $V$ , there is a function  $\lambda: V \rightarrow E \times V$ , that  $\lambda(v) = \{ \langle e_i, v_i \rangle | e_i \in E, v_i \in V \}$ . We call  $\lambda$  the semantic characteristic function,  $\{ \langle e_i, v_i \rangle \}$  the semantic pair set. This set is considered to be able to (uniquely ideally) identify  $v$ .



**Fig. 1** The contextual characteristics of a node  $N_i$  include its ontology or parent relationship  $f$  to its parent node  $N'$ , its inputting relationships  $I^*$  from the node  $N_{i-1}$ , its output  $O$  to the node  $N_{i+1}$ , and its attributes (i.e. metadata)  $M$ .

Our assumption is that if  $v$  is semantically related or semantically equivalent to  $w$ , where  $v$  and  $w$  are two nodes (concepts) in the graph  $G$ , then  $\lambda(v) \cap \lambda(w) \neq \emptyset$  or  $\lambda(v) = \lambda(w)$ , i.e. the set of contextual characteristics of  $v$  is intersected with or identical to that of  $w$ .

**Definition 2 (Semantic  $f$ -characteristic)** The semantic  $f$ -characteristic of  $v$ , denoted to be  $\lambda_f(v)$ , is a pair of edge  $e$  and node  $w$ , that is,  $\lambda_f(v) = \langle e, w \rangle$ . The generalized semantic  $f$ -characteristic of  $v$ , denoted to be  $\lambda_f^*(v)$ , is a set of pairs  $\langle e_i, w_i \rangle$ , where  $e_i$  and  $w_i$  are edges and nodes respectively and  $\lambda_f^*(w_{i-1}) = \langle e_i, w_i \rangle$ .

**Definition 3 (Semantic  $I$ -characteristic)** The semantic  $I$ -characteristic of  $v$ , denoted to be  $\lambda_I(v)$ , is a set of input and inputting node pairs, i.e.,  $\lambda_I(v) = \{ \langle i, w \rangle \mid i \in I \text{ and } w \in W \}$ . Here  $I$  and  $W$  are inputting relationships and nodes respectively in the graph  $G$ . We also define a function  $\#: 2^S \rightarrow P$  to indicate the size of the set, i.e.,  $\#(S) = p$ , where  $S \in 2^S$  (the power set of  $S$ ) and  $p \in P$  is a non-negative integer set. We also define one more metadata,  $T$ , called the type of an input edge. That is,  $T(i)$  is the type of the edge (relationship).

**Definition 4 (Semantic  $O$ -characteristic)** The semantic  $O$ -characteristic of  $v$ , denoted to be  $\lambda_O(v)$ , is a pair of output relationship and its connecting node, i.e.,  $\lambda_O(v) = \langle o, w \rangle$ , where  $o$  is the output edge from  $v$  to the node  $w$ . Like in the previous definition for  $I$  characteristic, we also define one more metadata,  $T$ , called the type of an output edge. That is,  $T(o)$  is the type of the edge (relationship).

### 3. Semantic Matches

A key step in construction of a virtual organization is the matching process where a sequence of tasks, which formally represents the users' demands, is semantically matched to a set of services, which are available in a service pool. The proposed contextual based semantic comparison and measure approach includes three steps: to compare the individual nodes forming the task sequence with the nodes from the

service pool, to find candidate service pairs that satisfy the task pairs, to assemble the service pairs into a sequence of services that meet the requirements of the sequence of tasks.

#### 3.1. Semantic relatedness for nodes

The first step to find a semantic match of two nodes is to compare their contextual characteristics of the nodes, i.e.  $f$ -characteristic,  $I$ -characteristic, and  $O$ -characteristic.

Suppose that  $v$  and  $w$  are two nodes (e.g. one task and one service) from two graphs,  $G$  and  $H$ . (Note that the symbol  $\equiv$  indicates that the two sets share some common elements.)

**Definition 5 (semantic  $f$ -related –  $ss-f$ )** We define that  $v$  and  $w$  are semantic  $f$ -related, i.e.  $(v, w) \in ss-f$  if  $\lambda_f^*(v) \equiv \lambda_f^*(w)$ .

**Definition 6 (semantic  $I$ -related –  $ss-I$ )** We define that  $v$  and  $w$  are semantic  $I$ -related, i.e.  $(v, w) \in ss-I$  if  $\lambda_I(v) \equiv \lambda_I(w)$  and  $\#(\lambda_I(v)) = \#(\lambda_I(w))$  and  $T(i_v) \equiv T(i_w)$ , where  $i_v \in I_v$  and  $i_w \in I_w$ .

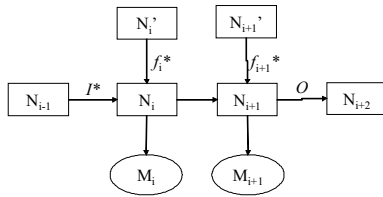
**Definition 7 (semantic  $O$ -related –  $ss-O$ )** We define that  $v$  and  $w$  are semantic  $O$ -related, i.e.  $\lambda_O(v, w) \in ss-O$  if  $\lambda_O(v) \equiv \lambda_O(w)$  and  $T(i_v) \equiv T(i_w)$ , where  $i_v \in I_v$  and  $i_w \in I_w$ .

#### 3.2. Semantic relatedness for pairs of nodes

The second step is semantic comparison and match for node pairs. Suppose that  $G$  and  $H$  are two graphs. We consider two adjacent nodes,  $t_i$  and  $t_{i+1}$ , from the graph  $G$ . After the first step, for  $t_i$  and  $t_{i+1}$ , we get two sets of nodes,  $S_i = \{s_{i1}, \dots, s_{mi}\}$  and  $S_{i+1} = \{s_{i+1,1}, \dots, s_{i+1,n_{i+1}}\}$ . All these nodes are from the graph  $H$ . We obtain a set of service pairs and the services in each pair possess an interdependent relationship in terms of their inputs and outputs. The interdependence structure for a service pair is illustrated in Fig. 2.

**Definition 8 (semantic pair-related –  $spr$ )** We define that, for a given pair of nodes, denoted to be  $t_i-t_{i+1}$ , any pair of nodes  $s_i-s_{i+1}$  where  $s_i \in S_i = \{s_{i1}, \dots, s_{mi}\}$  and  $s_{i+1} \in S_{i+1} = \{s_{i+1,1}, \dots, s_{i+1,n_{i+1}}\}$ , is semantic pair-related to  $t_i-t_{i+1}$ , i.e.  $(t_i-t_{i+1}, s_i-s_{i+1}) \in spr$ , if both  $(t_i, s_i)$  and  $(t_{i+1}, s_{i+1})$  belong to the semantic relatedness relations:  $ss-f$ ,  $ss-I$ , and  $ss-O$  at the same time, and  $\lambda_O(s_i) \subset \lambda_I(s_{i+1})$ .

Simply put it, this definition means that if there is a suitable input/output interdependence (or match) between two service nodes  $s_i$  and  $s_{i+1}$  and if there is a semantic inclusion relation between  $t_i$  and  $s_i$ , and  $t_{i+1}$  and  $s_{i+1}$ , respectively, the pair  $s_i-s_{i+1}$  is a suitable candidate pair matching the pair  $t_i-t_{i+1}$ .



**Fig. 2** A semantic contextual representation of the pair of nodes  $N_i$  and  $N_{i+1}$ .

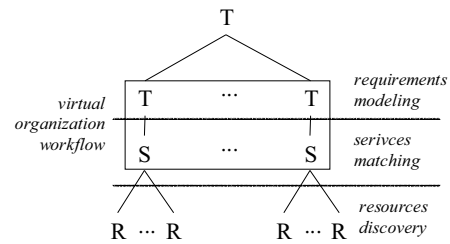
### 3.3. Semantic relatedness for node sequences

The last step is sequence semantic match, where we create a sequence of services out from the above-obtained candidate pairs of service nodes for the given sequence of tasks. Through iteratively applying the definitions given above, we get one or several candidate sequences of services that match the given task sequence. Here we should emphasize the importance of service scheduling which plays a critical role in this step because we consider that all the services used in the virtual organization are dynamically coupled and composed. Again, an important factor in the service scheduling process is how to measure quantitatively the semantic distance from the candidate services, service pairs, and service sequences to the given task sequence and its components (given tasks and task pairs). This approach has been further explored in [14].

### 3.4. Constructing workflow

Now, we see that there is a one-to-one mapping from the leaf tasks to the services. Using the preceding and the successor relations between the leaf tasks, we can get the following an ordered sequence:  $T_1 \ll T_2 \ll \dots \ll T_n$ , where  $T_i$ ,  $1 \leq i \leq n$ , is the  $i$ th leaf task in the leaf task set. We call this sequence as the workflow of this virtual organization and the corresponding sequence of the services,  $S_1 \ll S_2 \ll \dots \ll S_n$ , where  $S_i$ ,  $1 \leq i \leq n$ , is the  $i$ th service, as an implementation of the workflow. The formation of a virtual organization is illustrated in Fig. 3.

Here a virtual organization starts from a high level task for general users' requirements and demands. The task is decomposed into subtasks, which are in turn further decomposed, and hence form a task hierarchy. The leaf tasks form the workflow of the virtual organization. The task hierarchy consists of the modeling for the users' requirements. The services which match the leaf tasks form an implementation of the workflow. The service sequence layer with the workflow represents the service design and matching. The resources layer with the service sequence is the resource discovery and assembly.



**Fig. 3** The working structure of a virtual organization. Here T, S, and R stand for task, service, and resource. The leaf tasks together with the services form the virtual organization workflow.

## 4. Conclusion

Using the contextual semantic characteristics of a concept is an effective way to describe and identify the concept. In this paper, based on the work in [13], we proposed a contextual based semantic model to describe virtual organization components, in particular, tasks and services. One of the advantages is that the model is convenient to generate quantitative analysis on the components and hence easy to develop an automatic processing of semantic comparison and match of the components.

Virtual organizations have become an important subject in many application areas, for example e-Business, but how to semantically describe the components in a virtual organization and hence form a temporary and dynamic virtual organization is still a big problem as the diverse and distributed components, such as tasks, services, workflows, and resources, are very complex and volatile. Automation of virtual organization formation is our main goal in development of semantic web applications and semantic web service applications.

### \* Related Work and References

Due to the page limitation, the sections Related Work and References were omitted. A complete version can be obtained from <http://www.durham.ac.uk/w.w.song/wi2006-ws.pdf>.